File copy #3

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

SUBMARINE BASE, GROTON, CONN.

REPORT NUMBER 931

PERCEPTUAL CAPABILITIES REQUIRED TO OPERATE VISUAL SONAR DISPLAYS

by

Jo Ann S. Kinney
S. M. Luria

and

Alma P. Ryan

Naval Medical Research and Development Command
Research Work Unit M0100-PN.001-1002

Released by:

R. A. Margulies, CDR, MC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory
27 May 1980

PERCEPTUAL CAPABILITIES REQUIRED TO OPERATE VISUAL SONAR DISPLAYS

by

Jo Ann S. Kinney
S. M. Luria
Alma P. Ryan

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
REPORT NUMBER 931

Naval Medical Research and Development Command
Research Work Unit M0100-PN.001-1002

Released by:

R. A. Margulies, CDR, MC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

SUMMARY PAGE

PROBLEM

To determine which skills and abilities are required for effective operation of visual sonar displays.

FINDINGS

Success in operating visual sonar displays can be predicted moderately well by a few measures which include intellectual, perceptual, and experiential factors.

APPLICATION

The results could be used both to improve future manpower allocation and to suggest ways of improving operation of visual displays by current sonar technicians.

ADMINISTRATIVE INFORMATION

This investigation was conducted under Naval Medical Research and Development Command Work Unit M0100PN001-1002 --"Visual requirements for operators of visual sonar displays." It was submitted for review on 19 May 1980, approved for publication on 27 May 1980, and designated as NavSubMedRschLab Report No. 931 .

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

Abstract

The purpose of this research was to determine which skills and abilities are needed for operators of visual sonar displays. A standard procedure for occupational testing was employed: that is, a large group of men was given a battery of tests of abilities (visual, perceptual, cognitive, etc.) that were thought to be important to being a good operator of visual sonar displays. The men were then evaluated by their supervisors and by other sonar technicians (peers) for proficiency as sonar operators and the results from the test battery were correlated with these assessments. Several analyses showed that a small group of measures from the test battery differentiated among good and poor operators. For example, the use of GCT, time in service, near acuity, near lateral phoria, texture discrimination, and an attitude test gave a moderate ($R \cong .50$) and highly significant ($P < .01$) correlation with proficiency ratings. The implications are that a number of perceptual, motivational, and intellectual factors are important for operators of visual sonar displays and that good performance can be adequately predicted if these factors are properly tested.

INTRODUCTION

Submarine sonar systems have always been sophisticated and elaborate detection systems that require considerable skill and training on the part of the operator. In the past, extensive research has been devoted to testing sonarmen and to developing tests for their selection that would predict their efficiency as operators. Since the primary skills required were auditory, the selection of men was always on this basis; indeed, up until 1976, candidates for sonar school were screened on a test of memory for auditory pitch.

New sonar systems, like the BQQ5 and BQQ6, have been changed to present a visual display rather than relaying the information primarily to the auditory sense. These displays are extremely complex, and, like the auditory system, must require extensive skill and training for effective operation. The particular skills or abilities required, however, are largely unknown, as are the ranges of individual differences in performance and the effectiveness of training.

The goal of this research, therefore, was to determine the particular skills and abilities required to be an effective operator of visual sonar displays. Standard procedures have evolved in the field of occupational testing to address such a question.[1] These procedures include (1) analysis of the job to be performed, (2) administration of a trial battery of tests, chosen from the results of the job analysis, to a large group of subjects; (3) validation of the results from the test battery against an independent criterion of job success determined for each of the subjects, (4) selection of the final test battery, and (5) cross-validation against a new group of subjects.

Following these procedures, we have administered the tests to a group of 100 trained sonar technicians and reported the results of a factor analysis of the various measures.[2] The independent criterion of job success, ratings by supervisors and peers, has been obtained for each of the sonar technicians, and this report presents the results of the validation of the test battery against this criterion.

PROCEDURE

The Test Battery

A wide variety of perceptual and cognitive tests were included in the test battery. These were chosen on the basis of a review of the literature and the results of a questionnaire completed by 17 experienced sonar technicians. The battery included a number of visual tests which varied from simple measures of basic abilities to assessments of complex perceptual skills, as well as some intellectual, motivational or experiential measures. The complete test battery has been described in a previous report;[2] briefly, it consisted of measures of visual acuity at near and far distances, phoria, depth perception, accommodation, refractive error, and perceptual factors of speed of closure, flexibility of closure, perceptual speed, and spatial orientation. Two perceptual tests, designed on the basis of current theories of visual perception, were presented on cathode ray tubes; these were texture discrimination and contrast sensitivity. In addition, a number of non-perceptual measures were included: these were the General Classification Test (GCT) and an Arithmetic test (ARI), results from which were obtained from the men's records; their ages; length of time spent in service; and an Internal/External attitude test, which measures the extent to which an individual either feels in control of the things which happen to him or feels that events are beyond his control. Since a number of indices were available from some of these tests, the entire battery yielded 37 different scores by which to assess the men.

The Independent Criterion - Ratings of Proficiency

Several ratings of the sonar technicians' proficiency as operators of visual sonar displays were obtained for each man. The raters were the sonar chief from each boat, the sonar or weapons officer, the leading petty officers (LPOs), and all of the other sonar technicians (peers). In order to assure that the technicians had had sufficient experience operating the BQQ5, the ratings were obtained after the ship had been to sea and the sonar had been operated for at least six months.

The aim was to obtain a complete evaluation, consisting of measures from the test battery and ratings by all raters, on all sonar technicians on a given boat. During the time intervening between tests and ratings, some men were lost by transfer or retirement. The total sample, at the end of the study consisted of test results on 113 men. Of these, ratings were obtained for 101 men from eight different ships: these were the PHILADELPHIA, GROTON, OMAHA, NEW YORK CITY, PARGO, ARCHERFISH, MEMPHIS, and HAMMERHEAD. Twelve to thirteen sonar technicians were generally tested from each submarine.

The ratings obtained from those in a supervisory capacity differed from those obtained from the peers. The sonar chief, officers and LPOs were asked to rate all the men on a scale from zero to nine. A sample rating scale is provided in Appendix A. The instructions given to these raters were as follows:

> "You are being asked to rate the sonarmen on your ship for over-all ability as an operator of visual displays. (This does not include skills in maintenance or repair.) Simply make a check under each name at the position where he falls on the scale. Rate every man with respect to all others and also with respect to your general knowledge of sonar operators. After you have gone through the names once, look back over it to make sure there are no inconsistencies. Feel free to make any changes you wish.
>
> In any small group of men, such as the sonar crew, there is only a small chance of having some one really great or terribly poor. Generally there will be several men who are about average and perfectly acceptable. It is often easiest to start the rating with these average men and then ask yourself how much better or poorer the other men are."

Since it is well known that raters differ in the absolute level of ratings that they give, all rates were transformed to Z scores before being used in the final analysis.

Peers were asked simply to nominate in order of excellence those three individuals that they believed to be the best over-all operators of the BQQ5 on their ship. They were told to use all aspects of operating the sonar system, such as detection and classification, in their evaluation but not to include maintenance and repair. There were two reasons for the decision to ask peers for nominations of only good operators. Of primary importance was the consideration that the men would be more cooperative if asked to nominate their best operators than if asked who were the worst. Negative user-reaction can be a problem in peer assessments.[3] Second, comparisons of the validity of positive and negative peer nominations have indicated that the former are superior indices.[4,5]

The peer nominations were evaluated by assigning an arbitrary value of 3 to the man rated highest, 2 to the next best, and 1 to the next. The sum of all values achieved by each technician was used as the basic measure of his proficiency, according to his peers. This sum was then converted to the percent achieved, out of the total score possible, to equate for differences in the number of men rated on different ships.

Statistical Procedures

The major technique employed was a multiple correlation between various measures from the test battery and the criterion of proficiency as a sonar operator. Three different measures of proficiency were used. The first was an average of the ratings, in terms of Z scores, by three or four men in a supervisory capacity over the sonar operators on the submarine; these included the sonar chief, the sonar or weapons officer, and one or two LPOs. The second was the rating (Z scores) made by the sonar chief only; his evaluations were considered separately since it was assumed he would be most familiar with all the sonar technicians on his ship. The third was the peer evaluations.

Multiple correlations were done by computer, thus facilitating repeated trials of many different measures to find the highest and most significant correlations achieved by the least number of measures.

In addition, comparisons were made of the scores achieved on the various tests by the best and the poorest sonar technicians. Individuals were included in the best and poorest groups according to their ratings by their sonar chief or their nominations by their peers.

RESULTS

Comparison of Different Measures of Proficiency

The ratings made by sonar supervisors differed in some ways from the nominations by peers. Since supervisors rated all the sonar technicians, all of the rates on each ship were transformed into Z scores with a mean of zero and a standard deviation of 1.0; these scores were, of course, normally distributed. The peer nominations, for each ship, resulted in a distribution of scores varying from 100% (if a given technician was rated "best" by all other technicians on the ship) to zero (if a technician received no votes). These scores yielded a skewed distribution, since a few technicians usually received the majority of the votes and three to five men on each ship received none.

Figure 1 shows the distributions for a typical ship, on which 14 men were rated. Despite the differing techniques used to rate proficiency by the supervisors and by the peers, there was good agreement between the results. The five best and five poorest operators, according to the sonar chief, are shown in the hatched areas of Fig. 1; the five poorest all received no nominations from peers, while the five best included almost all those rated best by peers.

Another indication of the agreement is shown in Table I; this shows the various ranks achieved by the man rated best on each ship by his peers. Agreement between raters is just as good when only the poorer operators are considered. Thirty-four percent of the men received zero nominations from their peers; the average Z score of these men, as rated by supervisors was -0.175 and -0.84 by the sonar chief. These average Z scores are significantly worse than the mean of all the men at $p = <.005$.

In addition, rank-order correlations between the Z scores by the sonar chief and the percentage of nominations by peers, determined separately for each ship, ranged from .737 to .846 and averaged .773 for all eight submarines. For the ship depicted in Fig. 1, it was .774.

We thus conclude that there is substantial agreement on the ships as to who can and cannot operate the BQQ5 effectively. It is this opinion that we will attempt to predict with the best battery.

Table I. Ranks obtained by the sonar technicians rated "best" by peers on each ship

Submarine	Peer	Sonar Chief	All supervisors
PHILADELPHIA	1	1	1
MEMPHIS	1	3	1
GROTON	1	1	1
OMAHA	1	2.5	4
NEW YORK CITY	1	1	1
ARCHERFISH	1	4	5
PARGO	1	2.5	2
HAMMERHEAD	1	6	2

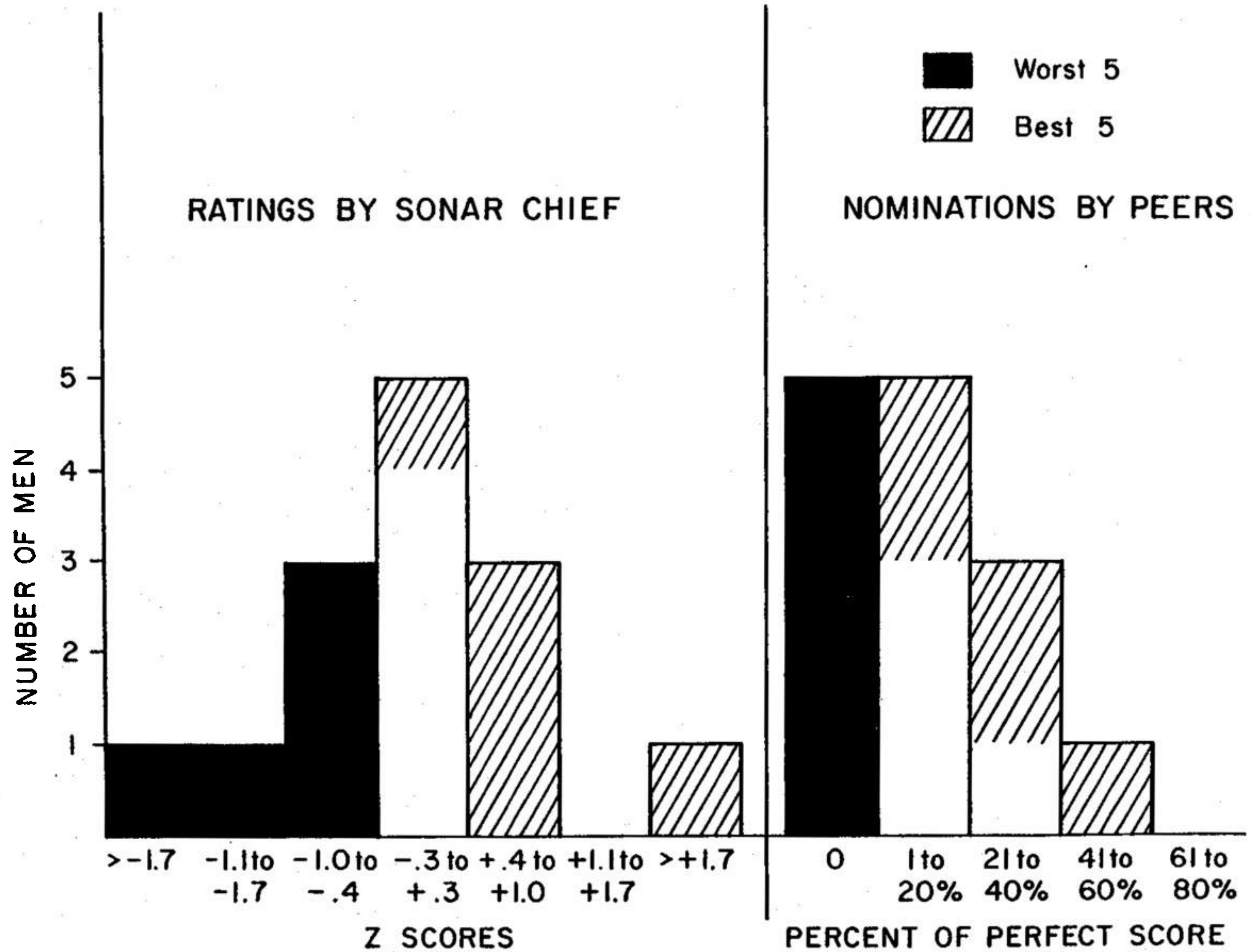

Fig. 1. The distribution of proficiency ratings for the 14 sonar technicians on a typical submarine. The shaded areas indicate the degree of agreement between ratings by the sonar chief and nominations by peers.

Multiple Correlations

Multiple correlations among the various 37 measures derived from the test battery were calculated to determine how well the ratings of job proficiency could be predicted. Three separate analyses were made, one for each of the three sets of ratings: the sonar chief's, the average ratings by 3 to 4 supervisors, and the peer ratings.

The aim of the analysis by multiple correlation was, of course, to obtain the highest and most significant predictive value, or R, from the least number of measures. At one extreme, no single test had a sufficiently high correlation with the performance measures to be an adequate predictor; at the other extreme, use of all 37 measures yielded the highest ($R = >.60$), but least significant predictor since many of the measures contributed nothing to the result.

Two procedures were helpful in selecting the best group of predictors. First, the results for the factor analysis, performed on the test battery,[2] proved useful. This analysis had yielded seven independent factors which best described the variance in the test battery. Theoretically, the use of more than one test from a single factor in a multiple correlation is counter-productive, since the tests in a single factor are correlated with one another and thus provide redundant information.[6] Empirically this proved to be true. The best example is Age and Time in Service, which are, of course, highly correlated. Prediction of job proficiencies from both measures resulted in an increase in the R of only .01 over the correlation with Time in Service alone. Similar results were obtained even with measures which have less intercorrelation, as long as they were from the same factor. For example, performance on the Hidden Patterns Test was correlated to some extent with job proficiency ($r = .19$). However, Hidden Patterns Test shares a common factor with the GCT and the use of the two in a multiple correlation resulted in an improvement of only .03 over the use of the GCT alone. When combined with several other important factors, this improvement drops to .005.

The second procedure was Thorndike's step-wise method[7] for determining the best R; in this analysis, one starts with the best predictor and adds one variable at a time until the addition of another adds but little. This step-wise procedure was followed, using only the best predictor from each factor, for each of the three sets of ratings of job proficiency.

Table II lists the six measures that added a significant increase to the R for one or more of the three criteria and the factor to which they belonged. These six measures were time in service, GCT, near acuity, near lateral phoria, and two different error scores from the texture test. The addition of the second texture is an exception to the general rule that the use of more than one measure from a single factor does not increase the multiple correlation. The exception occurs because of the unusual finding that for one of the measures, increasing errors were predictive of poor job proficiency, as might be expected, while for the other, increasing errors indicated good proficiency. The other two scores

Table II. Measures which contribute to multiple R's of Table III

Measure	Factor	$\underline{r}$ with criteria
Time in service	Time or age	.279
GCT	Paper & pencil tests	.217
Near acuity	Vision for fine detail	.178
Near lateral phoria	Horizontal eye alignment	-.154
Errors on discrimination of textures	Texture	.143
Errors on discrimination of textures	Texture	-.137
HITS	Vision for large objects	.114
Internal/External attitude test	none	-.127

included are HITS, the best measure of vision for large objects, and the Internal/External attitude test. This was included since it was the only one of the original measures which did not load (at the .3 level or higher) on any factor and is thus essentially a factor by itself. Of the original seven factors identified in the factor analysis, all are represented except the index of vertical eye alignment.

Table III gives the multiple correlation determined by the step-wise procedure for each of the three sets of criterion ratings and the corresponding F ratio. For example, the basic six measures correlated .495 with the Sonar Chief's ratings, and the F ratio was 5.08. The inclusion of the other two factors increased the correlation to .521 but reduced the F ratio to 4.28.

Two points are illustrated in these two tables. First, the six measures which contributed significantly to the multiple R do provide a moderately good prediction of the various criteria of job proficiency and these multiple correlations are substantially better than that provided by any single measure alone. Second, the addition of more measures produces only a small increment in the R while at the same time the over-all F ratio or level of significance falls slightly.

Table III. Multiple correlations determined by step-wise procedure

	Criterion Predicted		
	Sonar Chief's Ratings	Supervisor Ratings	Peer Nominations
6 measures which added significantly	.495	.532	.414
F ratio	5.08	6.17	3.24
8 measures 6 above plus two additional factors	.521	.546	.422
F ratio	4.28	4.89	2.49

Comparison of Good and Poor Operators

Another method of analyzing whether job proficiency can be predicted from the results of the test battery is a comparison of the average test scores of the best and poorest operators. For the assessment of job proficiency by the sonar chiefs, the 15 men who received the highest Z scores were compared with the 15 who received the lowest. For the peer assessment, a larger number were compared, since 32 of the 101 men received no nominations; these 32 constituted the group of poorer operators. Twenty-eight men received more than 20% of all possible nominations; these are the best operators.

Table IV shows the results of this comparison and lists all measures which differentiated between the groups at a probability of .10 or better. There is good agreement between the two measures of job proficiency and also between them and the results of the multiple correlations. The better operators tended to be older, had more time in service, higher GCTs, better near acuity, and were less exophoric.* They also made fewer errors on one of the texture discriminations and none on the other set, as noted above, and were more internal (that is, they believed that they, rather chance, are in control of the events in their lives). Only two discrepancies occur: ARI is included among the measures differentiating men according to the ratings of the sonar chiefs but not for the peers, while contrast thresholds for 10 cpd (an inverse measure of acuity) discriminates the groups according to the peers but not the sonar chiefs.

\- - - - - - - - - - - - - - -

* small numbers on this test indicate esophoria, large numbers, exophoria and the normal average is 7.5.

Table IV. Comparison of mean test results for the best and poorest sonar operators according to ratings by the sonar chiefs and peers.

Measures	Sonar Chiefs			Peers		
	Best operators N=15	Worst operators N=15	*t*	Best operators N=28	Worst operators N=32	*t*
Age	25.4	21.2	3.0	24.9	22.6	2.0
Time in service (mo)	56.5	15.9	5.8	58.8	33.0	2.3
GCT	61.7	57.7	2.0	62.0	59.3	1.8
ARI	58.0	52.1	1.4	57.8	58.9	-0.7
Internal/External attitude test	8.4	9.8	-0.9	7.7	9.1	-1.5
Near acuity (min. vis. angle)	1.2	1.1	0.9	1.2	1.1	1.3
Contrast thresholds for 10 cpd	0.016	0.016	0.0	0.016	0.021	-1.6
Phoria, near lateral	8.3	9.7	-1.2	7.9	9.0	-1.4
Texture (% errors on 52 patterns)	46.9	49.2	-1.1	47.1	49.0	-1.4
Texture (% errors on 17 patterns)	46.4	37.6	2.1	45.4	42.8	0.8

DISCUSSION

A common result in using a large number of predictors in a multiple correlation is that a few variables receive most of the β weights while the rest are near zero. Therefore, one of the critical issues in the use of multiple correlation is the determination of the smallest subset of predictors which accomplish most of the prediction. Thorndike lists three reasons for attempting to find the minimum number: economy, simplicity of description, and stability of the regression coefficient.[7] These were the reasons, of course, for employing the step-wise procedure to select the best tests.

Nonetheless, the selection of the specific tests to be employed always represents somewhat arbitrary decisions, since a number of measures make contributions of similar size to the final R, only one of which can be significant.* Since this is so, the results from all the analyses completed so far, the factor analysis, and the comparison of average data for good and poor operators, in addition to the multiple correlations, will be used in the final step of this research; that is, the cross-validation against a new sample of sonar operators. These measures that are considered to be important possible predictors are, then, time in service, GCT, near acuity, near lateral phoria, texture discrimination, and Internal/External attitude test.

The implications of these results seem reasonable and obvious. First, many factors are necessarily involved in the make-up of a good sonar operator on visual displays. Thus no one measure, of the 37 possible, would give an adequate prediction of job proficiency. Second, the measures that are required for adequate prediction come from many different areas--perceptual, intellectual, motivational, and experiential. Thus time in service is required in the correlation, showing that experience is an important factor. The Internal/External test provides a measure of the men's attitude on life and work and may be an essential ingredient. GCT is the best predictor from the whole group of tests of perceptual-cognitive skills, even though a GCT greater than one standard deviation above the mean is already a selection criterion for sonar technicians. Obviously modern submarines make sizeable intellectual demands on the operators. Finally, there are three visual or perceptual measures. While the texture discrimination may be an innate perceptual skill, near visual acuity and lateral phoria may reflect lack of proper optical corrections.

---

* For example, in predicting the peer ratings, the employment of the first six factors listed in Table III gives an R of .414; substituting the Internal/External test for the GCT, however, reduces the R only to .410, an obviously inconsequential difference.

REFERENCES

1. Anastasi, A. Psychological Testing (4th ed.) New York: Macmillan, 1976.

2. Kinney, J. A. S. and S. M. Luria. Factor analysis of perceptual and cognitive abilities tested by different methods. Perceptual & Motor Skills 50, 59-69, 1980.

3. Kane, J. S. and E. E. Lawler III. Methods of peer assessment. Psychological Bulletin 85(3) 555-586, 1978.

4. Kaufman, G. G. and J. C. Johnson. Scaling peer ratings: an examination of the differential validities of positive and negative nominations. Journal of Applied Psychology 59, 302-306, 1974.

5. Doll, R. E. and A. E. Longo. Improving the predictive effectiveness of peer ratings. Personnel Psychology 15, 215-220, 1962.

6. Kerlinger, F. N. and E. J. Pedhazur. Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston , 1973.

7. Thorndike, R. M. Correlational Procedures for Research. New York: Gardner Press Inc. 1978.

APPENDIX A

Men's names

Most competent I have ever known	9
	8
	7
Exceptionally competent	
	6
	5
Average competence	
	4
	3
Barely competent	
	2
	1
Completely incompetent	
	0

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NSMRL Report Number 931	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): PERCEPTUAL CAPABILITIES REQUIRED TO OPERATE VISUAL SONAR DISPLAYS	5. TYPE OF REPORT & PERIOD COVERED: Interim report
	6. PERFORMING ORG. REPORT NUMBER: NSMRL Rep. No. 931
7. AUTHOR(s): JO ANN S. KINNEY, S. M. LURIA and ALMA P. RYAN	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Submarine Medical Research Laboratory, Box 900, Naval Submarine Base, Groton, Connecticut 06340	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 65856N, M0100PN001-1002
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Submarine Medical Research Laboratory, Box 900, Naval Submarine Base, Groton, Connecticut 06340	12. REPORT DATE: 27 May 1980
	13. NUMBER OF PAGES: 11 plus appendix
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Naval Medical Research and Development Command, National Naval Medical Center, Bethesda, Maryland 20014	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

visual sonar displays; perceptual and cognitive tests; manpower allocation; sonar technicians

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The purpose of this research was to determine which skills and abilities are needed for operators of visual sonar displays. A standard procedure for occupational testing was employed: that is, a large group of men was given a battery of tests of abilities (visual, perceptual, cognitive, etc.) that were thought to be important to being a good operator of visual sonar displays. The men were then evaluated by their supervisors and by other sonar technicians (peers) for proficiency as sonar operators and the results from the test

(item 20--continued)

battery were correlated with these assessments. Several analyses showed that a small group of measures from the test battery differentiated among good and poor operators. For example, the use of GCT, time in service, near acuity, near lateral phoria, texture discrimination, and an attitude test gave a moderate ($R \cong .50$) and highly significant ($p < .01$) correlation with proficiency ratings. The implications are that a number of perceptual, motivational, and intellectual factors are important for operators of visual sonar displays and that good performance can be adequately predicted if these factors are properly tested.